34th Annual International Conference of the IEEE EMBS
San Diego, California USA, 28 August - 1 September, 2012

Automated Benign & Malignant Thyroid Lesion Characterization and Classification in 3D Contrast-Enhanced Ultrasound

U Rajendra Acharya, Vinitha Sree S*, Filippo Molinari, Roberto Garberoglio,

AgnieszkaWitkowska, Jasjit S Suri, *Fellow AIMBE*

***Abstract*—In this work, we present a Computer Aided Diagnosis (CAD) based technique for automatic classification of benign and malignant thyroid lesions in 3D contrast-enhanced ultrasound images. The images were obtained from 20 patients. Fine needle aspiration biopsy and histology confirmed malignancy. Discrete Wavelet Transform (DWT) and texture based features were extracted from the thyroid images. The resulting feature vectors were used to train and test three different classifiers: K-Nearest Neighbor (K-NN), Probabilistic Neural Network (PNN), and Decision Tree (DeTr) using ten-fold cross validation technique. Our results show that combination of DWT and texture features in the K-NN classifier resulted in a classification accuracy of 98.9%, a sensitivity of 98%, and a specificity of 99.8%. Thus, the preliminary results of the proposed technique show that it could be adapted as an adjunct tool that can give valuable second opinions to the doctors regarding the nature of the thyroid nodule. The technique is cost-effective, non-invasive, fast, completely automated and gives more objective and reproducible results compared to manual analysis of the ultrasound images. We however intend to establish the clinical applicability of this technique by evaluating it with more data in the future.**

***Index Terms*—Thyroid lesion, Computer Aided Diagnosis, Contrast Enhanced Ultrasound, Texture, Discrete Wavelet Transform**

I. Introduction

Thyroid nodules may occur in more than 50% of adult population. However, only about 7% of thyroid nodules are being diagnosed as malignant [1]. In order to effectively detect the small number of malignant nodules from the benign ones, there is a need for cost-effective accurate thyroid diagnosis support systems. Fine Needle Aspiration (FNA) biopsy can considered as the "gold standard" in the diagnosis of thyroid nodules [2]. However, FNA is too labor intensive to be used for large scale screenings and has pitfalls [3]. Computer Aided Diagnosis (CAD) based techniques are being developed to more objectively, automatically, quickly, and accurately analyze medical images and give valuable second opinions to doctors on the interpretations of the images. In this work, we have developed one such technique for thyroid nodule classification. High-Resolution Ultrasonography (HRUS) is cost-effective, commonly available, and is good for predicting the risk of malignancy [4]. Contrast-Enhanced Ultrasound (CEUS) imaging was introduced to enhance the differential diagnosis of solitary thyroid nodules [5]. It was found that CEUS enhancement patterns were different in benign and malignant lesions [6], and the application of ultrasonographic contrast agents led to an improvement in the differential diagnosis of thyroid nodules [7]. Due to the heavy vasculature of the thyroid gland, the contrast agent enhances the parenchyma representation, and presents the differences in the micro- and macro-vasculature in benign and malignant nodules. Thus, by using CEUS, it is possible to gain a better representation of the thyroid vascular pattern with a spatial resolution that is better than that of colour and power Doppler imaging. Therefore, in this study, we have used Discrete Wavelet Transform (DWT) and texture based features extracted from the CEUS images for developing the classification technique.

Manuscript received February 1, 2012.

U. Rajendra Acharya is with the Department of Electronics and Computer Engineering, Ngee Ann Polytechnic, Singapore

Vinitha Sree is with Global Biomedical Technologies Inc., CA, USA (email: vinitha.sree@gmail.com)

Filippo Molinari is with Biolab, Department of Electronics and Telecommunications, Politecnico di Torino, Torino, Italy.

Roberto Garberoglio is with Scientific Foundation Mauriziana Onlus, Torino, Italy

Jasjit S. Suri, Fellow AIMBE, is a CTO with Global Biomedical Technologies, CA, USA and is also affiliated with Biomedical Engineering Department, Idaho State University, ID, USA (jsuri@comcast.net).

II. Materials and Methods

Fig. 1 shows the block diagram of the proposed real-time CAD technique (a class of ThyroScan algorithms). It consists of an on-line classification system which predicts the class label (benign or malignant) of a test image based on the transformation of the on-line grayscale feature vector by the training parameters determined by an off-line learning system. The off-line classification system produces the training parameters using the combination of grayscale off-line training features and the respective off-line ground truth training class labels (0/1 for benign/malignant). The grayscale features for on-line or off-line training are based on image texture and Discrete Wavelet Transform (DWT). Significant features among the extracted ones are selected using the *t*-test. We evaluated K-Nearest Neighbor (K-NN), Probabilistic Neural Network (PNN), and Decision Tree (DeTr) classifiers. The above CAD system was developed using an image database, where the training set images were used to develop the classifiers. The built classifier was evaluated using the test set. For evaluation, we used a *k-fold* cross validation protocol. The predicted class labels of the test images and the corresponding ground truth labels (0/1) are compared to determine the performance measures of the

system such as sensitivity, specificity, accuracy, and Positive Predictive Value (PPV).

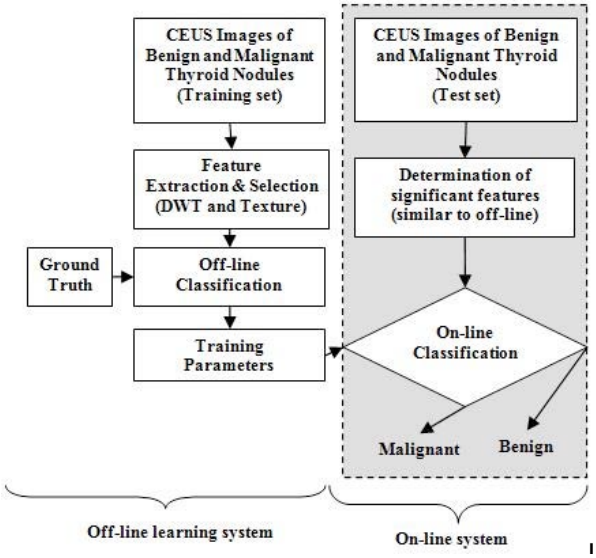

Fig. 1. Block diagram of the proposed system; The blocks outside the dotted shaded rectangular box represent the flow of off-line training system, and the blocks within the dotted box represent the on-line real-time system

A. Image Acquisition

Twenty patients (10 male and 10 female) with previously confirmed diagnosis of solitary thyroid nodule were enrolled in this study. All the patients signed an informed consent and the study was approved by the ethical committee of the "Umberto I'" Hospital of Torino. All subjects underwent a clinical examination, hormonal profiling, ultrasound (B-Mode and Color Doppler) examination, and FNA biopsy. Then, 2.5 ml of ultrasound contrast agent (Sonovue, Bracco, Italy) was administered intravenously and a 3D volume containing the lesion was acquired using MyLab70 ultrasound scanner (Biosound-Esaote, Genova, Italy) equipped by a LA-522 linear probe working in the range 4-10 MHz. All the images were acquired at 10 MHz. The average frame rate of the device was 16 Hz. Among the 20 nodules, five were benign (goiter nodules) and classified as THY2 according to cytological criteria [8]. Among the remaining 15 patients classified as THY3, five had benign conditions (follicular neoplasm) and ten had malignant conditions (seven papillary, one follicular and two Hurtle cells carcinoma). We acquired 40 data sets from each patient, and thus, we had 400 benign and 400 malignant images. The ten patients who were diagnosed with malignant nodules underwent thyroidectomy, and the histo-pathological analysis confirmed the diagnosis. The results of the FNAB were used as reference for the benign nodules. Fig. 2 and Fig. 3 show the typical malignant and benign thyroid images, respectively. The white arrows indicate the lesion. The red arrow indicates a portion of healthy parenchyma.

B. Feature Extraction

In the CEUS images, it was observed that ring enhancement was predominant in benign lesions and heterogeneous enhancement was prevalent in malignant lesions [6]. These differences manifest as changes in the image texture, and hence, we used texture and DWT based features in this work.

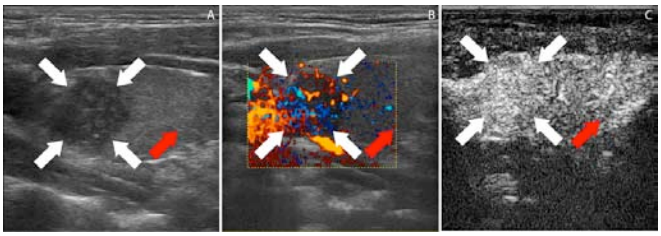

Fig. 2. A) HRUS, B) Color Doppler, C) CEUS images of a papillary carcinoma (malignant).

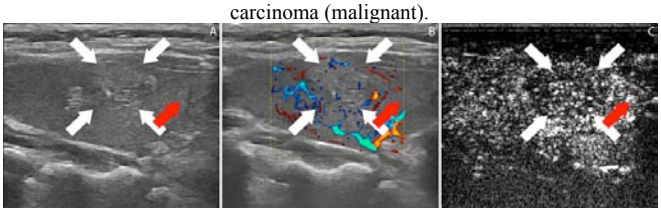

Fig. 3. A) HRUS, B) Color Doppler, C) CEUS images of a goiter nodule (benign).

DWT Features: DWT uses filter banks to decompose signals into low and high pass components, called sub-bands. The low pass coefficients contain information about slow varying characteristics, and the high pass components contain information about sudden changes. Low pass filtering (LPF) on both rows and columns results in LL coefficients which contain most of the image's total energy. LPF on rows and high pass filtering (HPF) on columns results in HL coefficients which contain the vertical details of the image. HPF on rows and LPF on columns results in LH coefficients which contain the horizontal details. Finally HPF on rows and columns results in the finest-scale HH coefficients that contain the diagonal details. Decomposition is further performed on the LL sub-band to attain the next coarser scale of wavelet coefficients. Fig. 4 shows the complete passband structure for a 2D sub-band transform with three levels of decomposition. A_2, V_2, H_2, D_2, H_1, V_1 and D_1 are the energies of LL_2, HL_2, LH_2, HH_2, LH_1, HL_1 and HH_1, respectively. In this work, we used Daubechies (Db) 8 as the mother wavelet. We stopped at the 2nd level decomposition, because higher level decompositions did not yield significant features. All the elements within the individual rows of the matrix were added and the elements of the resulting vector were squared before adding to form a scalar. This scalar was normalized by dividing it by the number of rows and columns of the original matrix to obtain the energy.

Texture Features: Texture features measure smoothness, coarseness, and regularity of pixels which form an image. In this work, we used the features extracted from the Gray Level Co-occurrence Matrix (GLCM). GLCM of an M × N image I is defined by

$$C_d(i,j) = \left| \begin{Bmatrix} (p,q),(p+\Delta x, q+\Delta y): I(p,q)=i, I \\ (p+\Delta x, q+\Delta y)=j \end{Bmatrix} \right| \quad (1)$$

where $(p,q),(p+\Delta x, q+\Delta y) \in M \times N$, $d=(\Delta x, \Delta y)$ and $|\cdot|$ denotes the cardinality of a set. The probability of a pixel with a grey level value i having a pixel with a gray level value j at a $(\Delta x, \Delta y)$ distance away in an image is

$$P_d(i,j)=\frac{C_d(i,j)}{\sum_i \sum_j C_d(i,j)} \quad (2)$$

Using Eqn. (2), we obtained the following features [9]:

$$Homogeneity \qquad H=\sum_i \sum_j \frac{P_d(i,j)}{1+(i-j)^2} \quad (3)$$

$$Entropy \qquad E=-\sum_i \sum_j P_d(i,j)\times \ln[P_d(i,j)] \quad (4)$$

$$Symmetry \qquad S=1-\sum_i \sum_j P_d(i,j)-P_d(j,i) \quad (5)$$

C. Classifiers

The **K-Nearest Neighbor (K-NN) classifier** is based on the minimum distance from a query instance to the training samples. The *K* nearest neighbors are determined using this method. A test sample is assigned the class that is the most common among its *K* nearest neighbors [10]. **Probabilistic Neural Network (PNN) classifier** is a type of two layer radial basis network in which the first layer of neurons has radial basis activation functions. This layer computes the distance vector by evaluating the distances between the input and training vectors. The second layer (competitive layer) sums the contributions of each input classes and produces a vector of probabilities as the output of the input classes. The compete transfer function at the output of the second layer selects the maximum of these probabilities and assigns a 1 for the selected class and a 0 for all other classes [11]. The **Decision Tree (DeTr) classifier** generates a tree and a set of rules to represent the model in order to identify different classes. The rules are used to predict the class of the test sample [12].

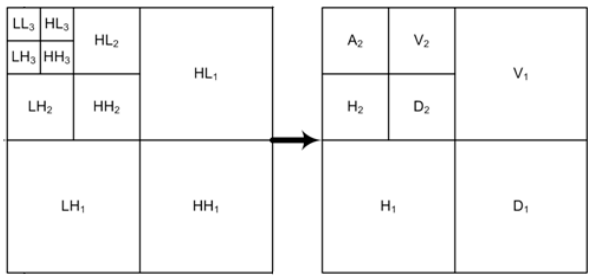

Fig. 4. Passband structure for a 2D sub-band DWT with three levels

III. Results

A. Selected Features

We used Student's *t-test* [13] to select the significant features. Table I documents the Mean ± Standard Deviation (SD) of these significant features for both benign and malignant classes. The fact that all p-values are below 0.0001 indicates that all features are clinically significant. The homogeneity, symmetry and all the DWT features are higher for malignant nodules compared to benign nodules because benign images (Fig. 3) have more structure in their sonographic appearance compared to malignant thyroid images (Fig. 2). The images with more structure, such as the benign thyroid images, have more variations in the greyscale values compared to the malignant thyroid images, and therefore have higher entropy values.

TABLE I
VALUES OF THE VARIOUS DWT AND TEXTURE FEATURES FOR THE BENIGN AND MALIGNANT THYROID LESIONS

Feature	Asymptomatic	Symptomatic	*p*-value
H	0.285 ±1.742E-02	0.345 ±4.280E-02	< 0.0001
E	3.78 ±5.402E-02	3.52 ±0.160	< 0.0001
S	0.851 ±1.355E-02	0.860 ±2.941E-02	< 0.0001
A_2	1.635E+06 ±6.189E+05	3.446E+06 ±2.335E+06	< 0.0001
H_2	4.20 ±2.30	5.12 ±3.06	< 0.0001
H_1	0.362 ±0.237	0.474 ±0.334	< 0.0001
V_2	319 ±124	676 ±665	< 0.0001
V_1	66.1 ±42.6	206 ±214	< 0.0001
D_2	1.07 ±0.626	2.20 ±1.20	< 0.0001
D_1	2.907E-02 ±6.303E-03	3.362E-02 ±1.744E-02	< 0.0001

B. Classification Results

Ten-fold stratified cross validation method was used to evaluate these classifiers. In this method, the dataset was split into ten parts, each part containing the same proportion of images from both classes. In the first fold, nine parts were used for training and the tenth part was used for testing and for calculation of the performance measures. This protocol was repeated nine more times with a different part as the test set. The averages of the performance measures (sensitivity, specificity, PPV, and accuracy) obtained during the testing phase of each fold are reported as the final performance measures for that classifier. Table II presents the classification results obtained by using the extracted DWT and texture features in the three classifiers. The classification accuracy for all three tested classifiers is well above 96%. It is also evident from the results that the K-NN classifier performs better than DeTr and PNN with a higher accuracy of 98.9%. It not only presents the highest accuracy but also has equally good sensitivity of 98% and specificity of 99.8%.

IV. Discussion

HRUS sensitivity and specificity were only 80.8% and 81.6%, respectively [14]. Color and Power Doppler cannot evidence micro vessels which are typical of malignancy [15] and also they exhibited poor specificity [16]. Finley *et al.* [17] classified benign and malignant thyroid nodules using

molecular profiling and obtained sensitivity and specificity of 91.7% and 96.2%, respectively. Patton et al. [18] obtained a low accuracy of 77% using fluorescent scanning. Based on the above studies, we felt the necessity for a better technique that can improve the classification efficiency and that is also more economical. In our previous study [19], we applied an advanced image processing and modeling technique to the CEUS 3D volumes of benign and malignant lesions, and quantified the parameters of the intranodal vascularity. In this paper, we developed a CAD classification framework for the differential diagnosis of thyroid lesions using 3D CEUS images that presents a high accuracy of over 98%. We believe that this high value of accuracy is due to the choice of the features used in the classifiers. DWT features extract the information from the images in both the time and frequency domains, and hence, these features capture the subtle variations in the characteristics of the benign and malignant images. In the case of texture features, the homogeneity feature measures the similarity between two pixels that are $(\Delta x, \Delta y)$ apart. Denseness and degree of disorder in an image are measured by the entropy feature. The symmetry projections indicate prominent directions within the texture of CEUS images, and therefore, symmetry is an important discriminative feature of these images. The proposed technique can be programmed into software which can be written in a CD and shipped to hospitals or be downloaded from the internet at no extra cost. On the limitations front, we have evaluated the technique using a small dataset. Since only about 7% of the thyroid nodules are generally malignant, building a large database of malignant nodules requires a big effort. We are currently working to enlarge our database.

TABLE II
CLASSIFICATION RESULTS; A: ACCURACY; SN: SENSITIVITY; SP: SPECIFICITY

Classifier	Accuracy (%)	PPV (%)	Sensitivity (%)	Specificity (%)
K-NN	98.9	98	99.8	98.9
PNN	97.8	95.8	99.8	97.8
DeTr	96.9	94	99.8	96.9

V. CONCLUSION

In this work, we investigated a CEUS based thyroid nodule classification CAD system that uses texture and DWT parameters. Our results show that the combination of DWT and texture features coupled with K-NN classifier presented a classification accuracy of 98.9%, sensitivity of 98% and 99.8% specificity. We believe that our strategy could be a step forward in the quantitative and user-independent classification of thyroid lesions. Since the cost associated with the CEUS examination followed by the use of the proposed CAD technique is relatively less than the cost of a surgical intervention, this CEUS based classification scheme can be considered as cost-effective. Moreover, the high accuracy encourages us to validate the system using a larger dataset in order to establish its clinical applicability as an adjunct tool to assist doctors in thyroid nodule classification and subsequent treatment regime.

REFERENCES

[1] C.D. Lansford and T.N. Teknos, "Evaluation of the Thyroid Nodule," *Cancer Control,* vol. 13, no. 2, pp. 89-98, 2006.

[2] Z. W. Baloch, S. Fleisher, V. A. LiVolsi, and P.K. Gupta, "Diagnosis of follicular neoplasm: a gray zone in thyroid fine-needle aspiration cytology," *Diagn. Cytopathol*. vol. 26, no. 1, pp. 1-44, Jan. 2002.

[3] N. P. Caraway, N. Sneige, and N. A. Samaan, "Diagnostic pitfalls in thyroid fine-needle aspiration: a review of 394 cases," *Diagn. Cytopathol.* vol. 9, no. 3, pp. 345-350, 1993.

[4] S. Bastin, M.J. Bolland, and M.S. Croxson, "Role of ultrasound in the assessment of nodular thyroid disease," *J. Med. Imaging Radiat. Oncol.,* vol. 53, no. 2, pp. 177-187, Apr. 2009.

[5] T. V. Bartolotta, M. Midiri, M. Galia, G. Runza, M. Attard, G. Savoia, R. Lagalla, and A. E. Cardinale, "Qualitative and quantitative evaluation of solitary thyroid nodules with contrast-enhanced ultrasound: initial results," *Eur. Radiol.*, vol. 16, no. 10, pp. 2234-2241, Oct. 2006.

[6] B. Zhang, Y. X. Jiang, J. B. Liu, M. Yang, Q. Dai, Q. L. Zhu, and P. Gao, "Utility of contrast-enhanced ultrasound for evaluation of thyroid nodules," *Thyroid*, vol. 20, no. 1, pp. 51–57, Jan. 2010.

[7] M. Appetecchia, D. Bacaro, R. Brigida, D. Milardi, A. Bianchi, and F. Solivetti, "Second generation ultrasonographic contrast agents in the diagnosis of neoplastic thyroid nodules," *J. Exp. Clin. Cancer Res.*, vol. 25, no. 3, pp. 325-330, Sep. 2006.

[8] R. Mihai, A. J. Parker, D. Roskell, and G. P. Sadler, "One in four patients with follicular thyroid cytology (THY3) has a thyroid carcinoma," Thyroid, vol. 19, no. 1, pp. 33-37, Jan. 2009.

[9] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Trans. Syst. Man. Cybern.* SMC-3, pp.610–621, 1973

[10] D. T. Larose, *Discovering Knowledge in Data: An introduction to data mining*, Chapter 5: KNN, New Jersey, USA: Wiley Interscience, 2004, pp. 90-106.

[11] D. F. Specht, "Probabilistic Neural Networks," *J. Neural Networks*, vol. 3, no. 1, pp. 109-118, 1990.

[12] D. T. Larose, *Discovering Knowledge in Data: An introduction to data mining*, Chapter 6: Decision Trees, New Jersey, USA: Wiley Interscience, 2004, pp. 108-126.

[13] J. F. Box, "Guinness, gosset, fisher, and small samples," *Statist. Sci*.,vol. 2, pp. 45–52, 1987.

[14] M. M. D'Souza, R. K. Marwaha, R. Sharma, A. Jaimini, S. Thomas, D. Singh, M. Jain, P. J. Bhalla, M. Tripathi, A. Tiwari, A. Mishra, A. Mondal, and R. P. Tripathi, "Prospective evaluation of solitary thyroid nodule on 18F-FDG PET/CT and high-resolution ultrasonography," *Ann. Nucl. Med.,* vol. 24, no. 5, pp. 345-355, Jun. 2010.

[15] G. Argalia, S. De Bernardis, D. Mariani, T. Abbattista, A. Taccaliti, L Ricciardelli, S. Faragona, P. M. Gusella, and G. M. Giuseppetti, "Ultrasonographic contrast agent: evaluation of time-intensity curves in the characterisation of solitary thyroid nodules," *Radiol. Med.* vol. 103, no. 4, pp. 407-413, Apr. 2002.

[16] M. C. Frates, C. B. Benson, P. M. Doubilet, E. S. Cibas, and E. Marqusee, "Can color Doppler sonography aid in the prediction of malignancy of thyroid nodules?," *J. Ultrasound Med.*, vol. 22,no. 2, pp. 127-131, Feb. 2003.

[17] D.J. Finley, B. Zhu, C.B. Barden, T.J. Fahey, "Discrimination of benign and malignant thyroid nodules by molecular profiling," *Ann. Surg.,* vol. 240, no. 3, pp. 425–437, Sep. 2004.

[18] J.A. Patton, J.W. Hollifield, A.B. Brill, G.S. Lee, D.D. Patton, "Differentiation between malignant and benign solitary thyroid nodules by fluorescent scanning," *J. Nucl. Med.*, vol. 17, no. 1, pp. 17–21, Jan. 1976.

[19] F. Molinari, A. Mantovani, M. Deandrea, P. Limone, R. Garberoglio, and J. S. Suri, "Characterization of single thyroid nodules by contrast-enhanced 3-D ultrasound," *Ultrasound Med. Biol.*, vol. 36, no. 10, pp. 1616-1625, Oct. 2010.